# Information Theory Exercise Sheet #2

University of Amsterdam, Master of Logic, Spring 2014
Lecturer: Christian Schaffner
TA: Hoang Cuong

Out: Thu, 13 February 2014
Due: Thu, 20 February 2014, 11:00

## To be solved in Class

1. *Disprove the teacher.* Let $n = \log(|\mathcal{X}|)$.
   (a) Give a joint distribution $P_{XY}$ where $H(X) = n$, and $X$ and $Y$ are dependent.
   (b) Prove that $H(X|Y) = n$ implies that $X$ and $Y$ are independent.

2. [Cover-Thomas 2.32]. We are given the following joint distribution of $X \in \{1, 2, 3\}$ and $Y \in \{a, b, c\}$:

$$P_{XY}(1, a) = P_{XY}(2, b) = P_{XY}(3, c) = 1/6$$
$$P_{XY}(1, b) = P_{XY}(1, c) = P_{XY}(2, a) = P_{XY}(2, c) = P_{XY}(3, a) = P_{XY}(3, b) = 1/12.$$

   Let $\hat{X}(Y)$ be an estimator for $X$ (based on $Y$) and let $p_e = P(\hat{X} \neq X)$.
   (a) Find an estimator $\hat{X}(Y)$ for which the probability of error $p_e$ is as small as possible.
   (b) Evaluate Fano's inequality for this problem and compare.

## Homework

1. [3 points] Show that the value

$$R(X; Y; Z) = I(X; Y) - I(X; Y|Z)$$

   is invariant under permutations of its arguments.

2. [6 points] Let $X, Y, Z$ be arbitrary random variables, and let $f$ be any deterministic function acting on $\mathcal{Y}$. In the following, replace "?" by "$\geq$" or "$\leq$" to obtain the correct inequalities, and reason each time with the help of an entropy diagram. **Hint:** $H(f(Y)|Y) = 0$.
   (a) $H(f(Y)) \,?\, H(Y)$
   (b) $H(X|f(Y)) \,?\, H(X|Y)$
   (c) $I(X; Z|Y) = 0$ implies $I(X; Z) \,?\, I(X; Y)$ and $I(X; Z) \,?\, I(Y; Z)$.

3. [6 points] For each statement below, specify a joint distribution $P_{XYZ}$ of random variables $X$, $Y$ and $Z$ ($P_{XY}$ of $X$ and $Y$ in (a)) such that the inequalities hold.
   (a) There exists a $y$, such that $H(X|Y = y) > H(X)$
   (b) $I(X; Y) > I(X; Y|Z)$
   (c) $I(X; Y) < I(X; Y|Z)$

**Note** that the distributions have to be different from the ones seen as examples during the lecture.

4. *Bottleneck.* Suppose a Markov chain starts in one of $n$ states, necks down to $k < n$ states, and then fans back to $m > k$ states. Thus $X_1 \to X_2 \to X_3$, i.e.,

$$P_{X_1X_2X_3}(x_1, x_2, x_3) = P_{X_1}(x_1) \cdot P_{X_2|X_1}(x_2|x_1) \cdot P_{X_3|X_2}(x_3|x_2)$$

for all $x_1 \in \{1, 2, ..., n\}$, $x_2 \in \{1, 2, ..., k\}$, $x_3 \in \{1, 2, ..., m\}$.

   (a) [3 points] Show that the dependence of $X_1$ and $X_3$ is limited by the bottleneck by proving that $I(X_1; X_3) \leq \log k$.

   (b) [1 point] Evaluate $I(X_1; X_3)$ for $k = 1$, and conclude that no dependence can survive such a bottleneck.

5. [4 points] *Conditional mutual information.* Consider a sequence of $n$ binary random variables $X_1, X_2, \ldots, X_n$. Each sequence with an even number of 1's has probability $2^{-(n-1)}$ and each sequence with an odd number of 1's has probability 0. Find the mutual informations

$$I(X_1; X_2), \quad I(X_2; X_3|X_1), \quad \ldots, \quad I(X_{n-1}; X_n|X_1, ..., X_{n-2}).$$

6. [6 points] *Run-length coding.* Let $X_1, X_2, \ldots, X_n$ be (possibly dependent) binary random variables. Suppose one calculates the run lengths $R = (R_1, R_2, \ldots)$ of this sequence (in order as they occur). For example, the sequence $X = 0001100100$ yields run lengths $R = (3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, \ldots, X_n)$, $H(R)$ and $H(X_n, R)$. Show all equalities and inequalities, and bound all the differences.